

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-046539
 (43)Date of publication of application : 14.02.2003

(51)Int.CI. H04L 12/46
 H04L 12/56

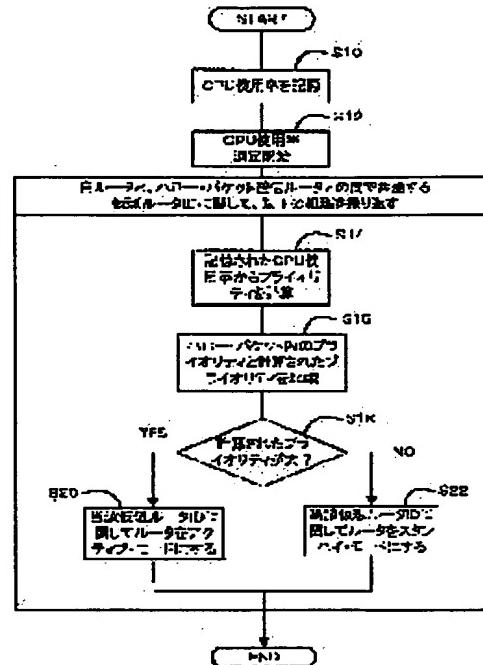
(21)Application number : 2001-233114 (71)Applicant : NEC CORP
 (22)Date of filing : 01.08.2001 (72)Inventor : SHIGEHASHI TAKAYUKI

(54) NETWORK RELAY DEVICE PROVIDED WITH LOAD DISTRIBUTING FUNCTION, LOAD DISTRIBUTING METHOD FOR NETWORK RELAY DEVICE AND PROGRAM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a load distribution method for a router, by which the processing load of the router is distributed.

SOLUTION: The router has a CPU operating rate measurement means, that measures a CPU operating rate of its own and a priority setting means that dynamically sets priority by using the CPU operating rate. The priority setting means is controlled so as to receive lower priority, when the CPU operating rate is higher; while in the case of a protocol, such as the VRRP(virtual router redundancy protocol), a router with the highest priority is assigned to particular packet processing among a plurality of the routers. Thus, a router group to which the priority, is set as stated above and the protocol above are combined, to enable the router whose CPU operating rate is low to process many packets, so as to assign the processing in a way of distributing the processing load among the routers.



LEGAL STATUS

[Date of request for examination] 19.07.2002

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

NOT AVAILABLE COPY

This Page Blank (uspto)

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-46539

(P2003-46539A)

(43) 公開日 平成15年2月14日 (2003.2.14)

(51) Int.Cl'	識別記号	F I	マーク(参考)
H 04 L 12/46	1 0 0	H 04 L 12/46	1 0 0 R 5 K 0 3 0
12/56	1 0 0	12/56	1 0 0 A 5 K 0 3 3

審査請求 有 請求項の数15 OL (全 11 頁)

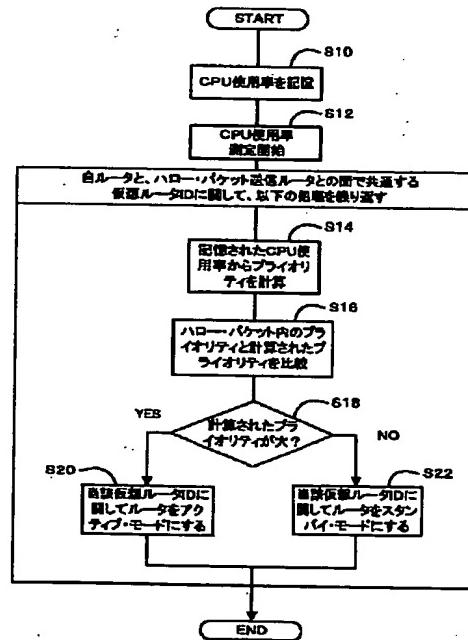
(21) 出願番号	特願2001-233114(P2001-233114)	(71) 出願人	000004237 日本電気株式会社 東京都港区芝五丁目7番1号
(22) 出願日	平成13年8月1日(2001.8.1)	(72) 発明者	重橋 敬之 東京都港区芝五丁目7番1号 日本電気株式会社内
		(74) 代理人	100071272 弁理士 後藤 洋介 (外1名) Fターム(参考) 5K030 GA11 HA08 HC14 HD03 HD08 JA11 JL07 KA05 LA03 LB05 LE03 LE05 MB02 MB09 MD02 5K033 AA03 CB06 CB18 CC01 DB18 EA03 EA07 EB06

(54) 【発明の名称】 負荷分散機能を有するネットワーク中継装置、ネットワーク中継装置の負荷分散方法、及びプログラム

(57) 【要約】 グラム

【課題】 ルータの処理負荷を分散させる、ルータの負荷分散方法が提供される。

【解決手段】 ルータは、自身のルータのCPU使用率を測定するCPU使用率測定手段と、そのCPU使用率を使用してプライオリティを動的に設定するプライオリティ設定手段を有する。プライオリティ設定手段は、CPU使用率が高いほど、低いプライオリティが設定されるように制御される。一方で、VRRP等のプロトコルでは、複数のルータにおいて、プライオリティの高いルータを、特定のパケットの処理に割り当てる。従って、上述のようにプライオリティが設定されたルータ群と、上記プロトコルを組み合わせることによって、絶えずCPU使用率の低いルータに多くのパケットを処理させて、ルータ間の処理負荷を分散させるよう割り当てが行われる。



【特許請求の範囲】

【請求項1】 ネットワーク中継装置において、プライオリティを記憶するプライオリティ記憶手段、前記プライオリティを、他のネットワーク中継装置のプライオリティと比較し、自身のプライオリティの方が他のネットワーク中継装置のプライオリティより大きい場合に、受信データの中継処理を行うよう制御するプライオリティ比較手段、前記ネットワーク中継装置の処理負荷を測定する負荷測定手段、及び前記負荷測定手段による測定結果を使用して、前記ネットワーク中継装置の負荷が高いほど低い値になるようにプライオリティを設定するプライオリティ設定手段とを有することを特徴とするネットワーク中継装置。

【請求項2】 請求項1に記載のネットワーク中継装置において、前記負荷測定手段が、前記ネットワーク中継装置の資源の使用頻度を測定することを特徴とするネットワーク中継装置。

【請求項3】 請求項2に記載のネットワーク中継装置において、前記負荷測定手段が、前記ネットワーク中継装置のCPU使用率を測定することを特徴とするネットワーク中継装置。

【請求項4】 請求項3に記載のネットワーク中継装置において、前記プライオリティ設定手段が、プライオリティPRI1を次式によって求めることを特徴とするネットワーク中継装置。

$$PRI1 = PRI1(B) - K1 * W/100$$

但し、PRI1(B)、K1は、事前に設定される固定値であり、Wは前記CPU使用率である。

【請求項5】 請求項1に記載のネットワーク中継装置において、前記負荷測定手段が、前記ネットワーク中継装置が所定時間内に受信する前記受信データの数を測定することを特徴とするネットワーク中継装置。

【請求項6】 請求項5に記載のネットワーク中継装置において、前記プライオリティ設定手段が、プライオリティPRI2を次式によって求めることを特徴とするネットワーク中継装置。

$$PRI2 = PRI2(B) - K2 * Y$$

但し、PRI2(B)、K2は、事前に設定される固定値であり、Yは前記受信データの数である。

【請求項7】 請求項1に記載のネットワーク中継装置において、前記ネットワーク中継装置がルータであり、前記ルータが、VRRPに基づいて前記プライオリティの比較を行い、かつ、受信データの中継処理を行うことを特徴とするネットワーク中継装置の負荷分散方法。

【請求項8】 請求項1に記載のネットワーク中継装置において、前記ネットワーク中継装置のそれぞれが仮想識別ID毎にプライオリティを有し、前記プライオリティの比較が同じ仮想識別IDの間で行われた結果、一のネットワーク中継装置において、他のネットワーク中継装置のプライオリティより大きいプライオリティを有する仮想識別IDが存在する場合に、前記一のネットワーク中継装置が、その仮想識別IDと同じIDに関連づけられたノードから送信された前記受信データの中継処理を行うよう制御することを特徴とするネットワーク中継装置。

【請求項9】 ネットワーク中継装置に記憶されたプライオリティを、他のネットワーク中継装置のプライオリティと比較するステップ、

前記ネットワーク中継装置のプライオリティの方が他のネットワーク中継装置のプライオリティより大きい場合に、受信データの中継処理を行うよう制御するステップ、

前記ネットワーク中継装置の処理負荷を測定するステップ、及び前記負荷測定ステップによる測定結果を使用して、前記ネットワーク中継装置の負荷が高いほど低い値になるようにプライオリティを設定するステップとを有することを特徴とするネットワーク中継装置の負荷分散方法。

【請求項10】 請求項9に記載のネットワーク中継装置の負荷分散方法において、

前記負荷測定ステップが、前記ネットワーク中継装置の資源の使用頻度を測定することを特徴とするネットワーク中継装置の負荷分散方法。

【請求項11】 請求項10に記載のネットワーク中継装置の負荷分散方法において、前記負荷測定ステップが、前記ネットワーク中継装置のCPU使用率を測定することを特徴とするネットワーク中継装置の負荷分散方法。

【請求項12】 請求項9に記載のネットワーク中継装置の負荷分散方法において、前記負荷測定ステップが、前記ネットワーク中継装置が所定時間内に受信する前記受信データの数を測定することを特徴とするネットワーク中継装置の負荷分散方法。

【請求項13】 請求項9に記載のネットワーク中継装置の負荷分散方法において、前記ネットワーク中継装置がルータであり、前記ルータが、VRRPに基づいて前記プライオリティの比較を行い、かつ、受信データの中継処理を行うことを特徴とするネットワーク中継装置の負荷分散方法。

【請求項14】 コンピュータを、
プライオリティを記憶するプライオリティ記憶手段と、
前記プライオリティを、他のネットワーク中継装置のブ

ライオリティと比較し、自身のプライオリティの方が他のネットワーク中継装置のプライオリティより大きい場合に、受信データの中継処理を行うよう制御するプライオリティ比較手段と、

前記ネットワーク中継装置の処理負荷を測定する負荷測定手段と、

前記負荷測定手段による測定結果を使用して、前記ネットワーク中継装置の負荷が高いほど低い値になるようにプライオリティを設定するプライオリティ設定手段とを有するネットワーク中継装置として機能させるプログラム。

【請求項15】コンピュータを、
ネットワーク中継装置の処理負荷を測定する負荷測定手段と、

前記負荷測定手段による測定結果を使用して、前記ネットワーク中継装置の負荷が高いほど低い値になるようにプライオリティを設定するプライオリティ設定手段とを有するネットワーク中継装置として機能させるプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、VRRP (Virtual Router Redundancy Protocol) を利用したルータにおいて、負荷分散を可能とし、特に、VRRPにおいて使用される各ルータのプライオリティを動的に変化させることによってルータ間の負荷を分散させる方法に関する。

【0002】

【従来の技術】VRRPは、IP (インターネット・プロトコル) の1つであり、RFC2338に準拠し、LAN上で静的構成のルータを使う場合に、バックアップルータの存在を許すものである。

【0003】目的のノードやネットワークにパケットを送信するに複数の経路があって、送信元のノードが接続されるLAN上に、その複数の経路に対応する複数のルータが存在する場合、VRRPは、そのうちの1つのルータを選択し、送信元のノードからのパケットは、そのルータを経由して目的のノード等に送信される。

【0004】これらの複数のルータは、1つの仮想ルータ (Virtual Router) を形成し、送信元のノードからは、1つのIPアドレスを有するルータとして取り扱われる。従って、VRRPがどのルータを選んでも、送信元のノードからのパケットの送信は、1つのIPアドレスに対して行われる。

【0005】VRRPでは、各ルータ間において一定間隔でハロー・パケットを交換し、各ルータが正常な状態かどうかをチェックしており、これはヘルスチェックと呼ばれる。また、当該ハロー・パケットには、ハロー・パケットを送信したルータのプライオリティが含まれており、各ルータは、これらのプライオリティと自身のプライオ

リティを比較することによって、どのルータがパケットを処理すべきアクティブ・ルータ (マスタ・ルータ) かを判断する。即ち、正常で最もプライオリティの高いルータを自律的にアクティブ・ルータとして設定し、その他のルータをスタンバイ・ルータ (バックアップ・ルータ) とする。

【0006】マスタ・ルータとして設定されたルータは、1つの仮想ルータとして定義された1群のルータを代表して、受信したパケットを処理する。マスタ・ルータが障害等の理由で通信不能となった場合、他のバックアップ・ルータは、マスタ・ルータがハロー・パケットに応答しないことを検知し、これらのバックアップ・ルータの中で、プライオリティの最も高いルータを、前のマスタ・ルータと同じIPアドレスを有する次のマスタ・ルータに設定する。

【0007】こうしたVRRPの機能により、IPアドレスを変更することなく、通信不能となったルータを自動的に通信可能なルータに切り替えることができる。このことによって、1つのルータが障害となっても、パケットの送信を継続することができ、結果的に、ネットワークの信頼性が向上する。

【0008】

【発明が解決しようとする課題】上述のように、VRRPは、ルータの冗長機能を提供するプロトコルであるが、このプロトコルには2つの問題点がある。

【0009】第1の問題点は、マスタ・ルータが完全に通信不能にならないとルータの切替が行われないことがある。マスタ・ルータにトラフィックが集中して、パケットの処理効率が著しく低下している場合であっても、他のバックアップ・ルータに切り替わることがないため、特定のルータに処理が集中する傾向にある。

【0010】第2の問題点は、VRRPに負荷分散機能がないことである。これは、第1の問題点とも密接に関連する問題であるが、第1の問題点で述べたように、ルータの切り替えは、マスタ・ルータが完全に通信不能になったことによって行われるために、マスタ・ルータが稼働している間に、その負荷を判定してルータの切り替えを行い、他のルータに負荷を分散させるような制御はされていない。

【0011】

【課題を解決するための手段】従って、本発明は、VRRPを使用したルータにおいて、ルータ自身の負荷を判定し、負荷の低いルータがマスタ・ルータとなるようにルータの切り替えを行い、全体としてルータ間の負荷を分散させることができるルータ、負荷分散方法、及び当該方法を実行するプログラムを提供する。

【0012】本発明の第1の実施態様によれば、プライオリティを記憶するプライオリティ記憶手段、前記プライオリティを、他のネットワーク中継装置のプライオリティと比較し、自身のプライオリティの方が他のネット

ワーク中継装置のプライオリティより大きい場合に、受信データの中継処理を行うよう制御するプライオリティ比較手段、前記ネットワーク中継装置の処理負荷を測定する負荷測定手段、及び前記負荷測定手段による測定結果を使用して、前記ネットワーク中継装置の負荷が高いほど低い値になるようにプライオリティを設定するプライオリティ設定手段とを有するネットワーク中継装置が提供される。

【0013】これによって、ネットワーク中継装置の負荷が高いほど、そのネットワーク中継装置が処理を担当するデータは少なくなるように自動的に制御され、ネットワーク中継装置間で負荷が分散される。

【0014】本発明の第2の実施態様によれば、前記第1の実施態様のネットワーク中継装置において、前記負荷測定手段が、前記ネットワーク中継装置のCPU使用率を測定するネットワーク中継装置が提供される。

【0015】これによって、ネットワーク中継装置の負荷が、CPU使用率をもとに把握され、複数のネットワーク中継装置の間で、受信データの処理に関する負荷が分散される。

【0016】本発明の第3の実施態様によれば、前記第1の実施態様におけるネットワーク中継装置がルータであり、前記ルータが、VRRPに基づいて前記プライオリティの比較を行い、かつ、受信データの中継処理を行うように構成される。

【0017】これによって、各ルータは、自動的にルータ間で負荷を分散しながら、IPパケットの中継処理を行う。

【0018】本発明の第4の実施態様によれば、ネットワーク中継装置に記憶されたプライオリティを、他のネットワーク中継装置のプライオリティと比較するステップ、前記ネットワーク中継装置のプライオリティの方が他のネットワーク中継装置のプライオリティより大きい場合に、受信データの中継処理を行うよう制御するステップ、前記ネットワーク中継装置の処理負荷を測定するステップ、及び前記負荷測定ステップによる測定結果を使用して、前記ネットワーク中継装置の負荷が高いほど低い値になるようにプライオリティを設定するステップとを有するネットワーク中継装置の負荷分散方法が提供される。

【0019】これによって、ネットワーク中継装置の負荷が高いほど、そのネットワーク中継装置が処理を担当するデータは少くなるように自動的に制御され、ネットワーク中継装置間で負荷が分散される。

【0020】本発明の第5の実施態様によれば、プライオリティを記憶するプライオリティ記憶手段と、前記プライオリティを、他のネットワーク中継装置のプライオリティと比較し、自身のプライオリティの方が他のネットワーク中継装置のプライオリティより大きい場合に、受信データの中継処理を行うよう制御するプライオリテ

イ比較手段と、前記ネットワーク中継装置の処理負荷を測定する負荷測定手段と、前記負荷測定手段による測定結果を使用して、前記ネットワーク中継装置の負荷が高いほど低い値になるようにプライオリティを設定するプライオリティ設定手段とを有するネットワーク中継装置としてコンピュータを機能させるプログラムが提供される。

【0021】これによって、ネットワーク中継装置の負荷が高いほど、そのネットワーク中継装置が処理を担当するデータは少なくなるように自動的に制御され、ネットワーク中継装置間で負荷が分散される。

【0022】

【発明の実施の形態】最初に、図6を参照して、従来のVRRPの機能について、特に、グループ化されたプライオリティの取り扱いについて詳細に説明する。

【0023】図6には2つのルータ(71、72)と4つのホスト(81、82、83、84)がネットワークを介して接続されている。図には示されていないが、ルータ71、72は、更に別のルータ、又はネットワークに接続されており、ホストから受信したパケットのルーティングを行い、そのパケットを、決定された適当なルータ、又はネットワークに送信する。

【0024】図中、各ホストは、VRID(1)又はVRID(2)という情報を有しているが、これは、各ホストが、通常ルータとして用いるルータ(デフォルト・ルータ)を識別するための仮想ルータIDである。ホスト81とホスト82は、VRID(1)という仮想ルータIDを有し、ホスト83とホスト84は、VRID(2)という仮想ルータIDを有している。

【0025】一方、ルータ71、72は、それぞれ前記仮想ルータID毎に、プライオリティを有している。ルータ71は、VRID(1)についてPRI(1-1)を、VRID(2)についてPRI(1-2)を有し、ルータ72は、VRID(1)についてプライオリティPRI(2-1)を、VRID(2)についてPRI(2-1)を有している。

【0026】ここで、PRI(1-1)ないしPRI(2-2)は、互いに比較可能な所定の固定値を表すものである。

【0027】従って、この場合、VRRPは、ルータ71とルータ72の間でハローパケットにより上記プライオリティの値を交換して比較し、他のルータより値が大きければ、そのルータが、対応する仮想ルータIDに関してマスタ・ルータとして機能する。

【0028】例えば、VRID(1)について、PRI(1-1) > PRI(2-1)であれば、仮想ルータID=VRID(1)に関してマスタ・ルータとして機能するルータは、ルータ71ということになる。また、VRID(2)について、PRI(1-2) < PRI(2-2)であれば、仮想ルータID=VRID(2)に

してマスタ・ルータとして機能するルータは、ルータ72になる。

【0029】このように仮想ルータIDのプライオリティが判定されると、前述のように、ホスト81とホスト82は、仮想ルータID=VRID(1)をデフォルト・ルータとするので、これらのホストからのパケットは、ルータ71を経由して送信先に送信される。一方、ホスト83とホスト84は、仮想ルータID=VRID(2)をデフォルト・ルータとするので、これらのホストからのパケットは、ルータ72を経由することになる。この状態では、2つのホストからのパケットが、1つのルータを経由するように構成され、ある程度、ルータの負荷分散がされているということができる。

【0030】ルータ71とルータ72の間では、一定の間隔でハローパケットが交換され、そのタイミングで両者のプライオリティが比較される。しかし、前述通り、PRI(1-1)ないしPRI(2-2)は固定値であるため、これらの値が管理者等によって変更されない限り、上記ホストとルータの対応関係は変更されない。

【0031】ここで、ルータ71に障害が発生し、完全に機能が停止したと仮定すると、ルータ72は、自身が発したハローパケットの応答がルータ71からないとために、ルータ71が通信不能になったと判断し、それまでルータ71が行ってきた仮想ルータID=VRID

(1)としての機能を引き継ぎ、結果的に、2つの仮想ルータID(VRID(1)、及びVRID(2))のマスタ・ルータとして動作する。即ち、ホスト81ないしホスト84は全て、ルータ72を介してパケットを送信することになる。

【0032】この後、ルータ71が復旧して、元のプライオリティで立ち上がりると、ルータ72がルータ71に送信したハローパケットの応答がルータ71から返され、これによって、再びプライオリティの比較が行われ、ルータ71が再び仮想ルータID=VRID(1)

$$PRI1 = PRI1(B) -$$

ここで、PRI1(B)(基本プライオリティ)は、予め管理者又は利用者が設定する値であり、PRI1(H)(補正プライオリティ)は、CPU使用率Wを元

$$PRI1(H) = K1 * W / 100$$

但し、K1は所定の係数である。更に、Wは0から100(%)であるため、W/100は、0から1の間の値を有する。

【0041】また、CPU使用率は、それ自体、所定の時間当たりの利用割合を表すものであるが、ここでは、ある一定時間Tの間における、これらの値の最大値、最小値、平均値等が選択的に使用され得る。もちろん、これ以外の基準や、組み合わせによって、CPU使用率を表すことも可能である。

【0042】また更に、基本プライオリティPRI1

のマスタ・ルータとして機能し始める。

【0033】しかし、このような構成では、ルータ71及びルータ72の間の負荷の偏りを自動的に是正することはできない。例えば、ホスト81、及びホスト82のパケットがルータ71を経由して送信され、これらのホストからのパケット数がある時点で、他のホストのパケット数に比べて極端に増加し、ルータ71の処理能力を超えることになっても、VRRPではルータ71が完全に通信不能とならない限り、その処理の一部をルータ72が分担するような制御は行われない。

【0034】次に、図1を参照して、本発明の第1の実施形態のルータについて説明する。

【0035】前述のように、従来のVRRPにおいては、プライオリティは管理者等が設定した固定的な値であったが、本実施形態では、このプライオリティの値がルータのCPU使用率等の測定値に依存して動的に変化する。

【0036】図1のルータは、ルーティング・テーブルを使用したルーティング手段等の一般的手段の他に、CPU使用率測定手段11、及びプライオリティ設定手段12を有している。

【0037】CPU使用率測定手段11は、自身のルータにおけるCPU使用率を測定し、その結果を0から100(%)の間の値Wで返す機能を有する。CPU使用率は、通常は、ホストからのパケット数が増加すれば必然的に増加するものであるが、パケットの宛先が多岐にわたって、ルーティング・テーブルを頻繁に参照する必要がある場合など、受信するパケットの性質にも依存する。

【0038】プライオリティ設定手段12は、CPU使用率測定手段11で測定されたCPU使用率Wを元にして、動的にプライオリティを設定する機能を有する。プライオリティPRI1は、以下の式1によって求められる。

【0039】

$$PRI1(H) \dots \text{(式1)}$$

にして計算された値であり、以下の式2によって求められる。

【0040】

$$(B) \dots \text{(式2)}$$

(B)、係数K1は、ルータ毎に、そのルータの性能等に応じて設定することができ、また、1つのルータの中でも、仮想ルータID毎に異なる値を設定することができる。

【0043】本発明の第1の実施形態では、ルータの負荷分散を行うために、ルータのCPU使用率を測定し、その使用率を下げる方向に、即ち、CPU使用率の低いルータが、より高いプライオリティを有して、より多いパケットのルーティングを受け持つようにして、ルータ間の負荷を分散させようとするものであるが、ルータの

他の資源の使用頻度、例えば、メモリーの使用率やハードディスクのI/Oデータ量などをもとにしてプライオリティを計算することもできる。

【0044】また、こうしたルータの各資源の使用頻度をいくつか組み合わせてプライオリティを計算することも可能である。

【0045】プライオリティをルータ間で比較するタイミングは、VRRPにおいてハローパケットを受信するタイミング（通常、1秒間隔）が望ましい。従って、CPU使用率を測定する一定時間Tも、そのタイミング（即ち1秒）とすることが可能である。

【0046】しかし、ルータの性能によっては、このように頻繁にプライオリティを比較してマスタ・ルータの切り替えを行うことが却って全体の性能を劣化させる場合も考えられるので、上記測定タイミングは慎重に調整する必要がある。

【0047】また、CPU使用率を測定する一定時間Tを2秒、5秒等、ある程度長時間に設定し、CPU使用率の長期の変動傾向をプライオリティに反映させることもできる。

【0048】また更に、上記式1、及び式2は、CPU使用率の高いものほどプライオリティが小さくなるように調整された式の一例を示したに過ぎず、本発明の範囲がこれに限られるものではない。当業者であれば、他の多くのパターンを導出することが可能である。

【0049】次に、図1、及び図6を参照して、本発明の第1の実施形態のルータの動作について具体的に説明する。

【0050】定常状態において、PRI(1-1) > PRI(2-1)の関係が成立している場合、仮想ルータID=VRID(1)に関しては、ルータ71がマスタ・ルータとなり、ホスト81、及びホスト82のデフォルト・ルータとしてルータ71が選択される。

【0051】同様に、PRI(1-2) < PRI(2-2)の関係が成立しているとすると、仮想ルータID=VRID(2)においては、ルータ72がマスタ・ルータとなり、ホスト83、及びホスト84のデフォルト・ルータとしてルータ72が選択される。

【0052】この状態で、各ホストがパケットを必要に応じてそれぞれ対応するデフォルト・ルータに送信した結果、ルータ71が高負荷状態となり、一定時間Tの間、CPU使用率がWとなったとする（ここでは、一定時間TにおけるCPU使用率の最小値をWとする）。

【0053】前述した式1、式2により、PRI(1-1)を計算した結果、プライオリティの値が減少して、PRI(1-1) < PRI(2-1)の関係になると、仮想ルータID=VRID(1)に関しては、ルータ72がマスタ・ルータとなり、ホスト81、及びホスト82からのパケットは、ルータ72を経由して送信される

ことになる。

【0054】この結果、ルータ71の高負荷状態を回避することができる。また、このことによって、所定時間経過後の一定時間Tにおいて、ルータ71のCPU使用率が低下すると同時に、ルータ72のCPU使用率が上昇した結果、再びPRI(1-1) > PRI(2-1)の関係になると、ホスト81、及びホスト82のデフォルト・ルータは、再びルータ71に設定される。

【0055】こうしたVRRPの従来の機能と、プライオリティの動的な制御によって、負荷の高いルータの処理負担を他のルータに動的に分散して、全体として効率のよいパケット処理が実現される。

【0056】次に、図2、及び図3を参照して、本発明の第1の実施形態のルータにおける処理フローを説明する。

【0057】図2は、ある1つのルータが、他のルータからハローパケットを受信した際に起動されるプロセスのフローを表している。当該プロセスは、受信したハローパケット内に、相手ルータのプライオリティの値が含まれている関係上、このタイミングで起動されることが望ましいが、これ以外のタイミングで起動することも可能である。

【0058】また、ここでは、CPU使用率を測定する一定時間Tが、ハローパケットを受信するタイミング（通常、1秒）と同じであるとして説明する。もちろん、複数の記憶エリアを割り当てるによって、これ以上の時間におけるCPU使用率を何種類か並行的に記憶することができる。

【0059】ハローパケットを他のルータ（例えば、ルータ72）から受け取ると、ルータ（例えば、ルータ71）は、まず、ステップS10において、これまで測定していたCPU使用率をセーブし、ステップS12において、次のハローパケット受信時に使用する、新たな次の時間TのCPU使用率の測定を開始する。

【0060】その後、ステップS14ないしステップS22の処理が、ルータと他のルータとの間で共通する仮想ルータID毎に、繰り返し行われる。プライオリティの計算やプライオリティの比較が仮想ルータID毎にされるためである。

【0061】ステップS14では、ステップS10でセーブされたCPU使用率から、前述の式1、式2を使用してプライオリティを計算する。ステップS10でセーブされたCPU使用率は、前回のハローパケット受信から今回のハローパケット受信まで（時間T）のCPU使用率である。

【0062】次に、ステップS16では、ハローパケット内の、他のルータのプライオリティと、ステップS14で計算されたプライオリティを比較する。計算された自ルータのプライオリティの方が大きい場合（ステップS18のYES）、自ルータを、その仮想ルータIDに

関してアクティブ・モードにする（ステップS20）。これによって、その仮想ルータIDを有するホストからのパケットがそのルータによって取り扱われ、適切な経路で送信される。

【0063】それ以外の場合（ステップS18のNO）、自ルータを、その仮想ルータIDに関してスタンバイ・モードにする（ステップS22）。この場合、その仮想ルータIDのホストから送信されたパケットは、そのルータによって扱われない。

【0064】このフローが、各ルータで実行されるため、例えば、1つのルータ71が、他のルータ72からハローパケットを受信した際に、ある仮想ルータIDに関してスタンバイ・モードを選択したとすると、逆に、他のルータ72は、ルータ71からハローパケットを受信した際に、その仮想ルータIDに関してアクティブ・モードを選択することになる。これは、ルータ72の方が、その時点でのCPU使用率が低かったことを示しており、上記仮想ルータIDに関しての処理が、ルータ71の担当からルータ72の担当に振り向かれたことになる。こうして、ルータ間の負荷分散が自動的に行われ、全体として効率のよいパケット処理が行われる。

【0065】また、この例では、理解を容易にするため、2つのルータ、及び4つのホスト（ノード）を用いて説明がされているが、これらの数に制限はない。更に、仮想ルータIDの設定方法についても各ルータ間で適宜行うことができるが、冗長性を維持するため、1つの仮想ルータIDは、通常複数のルータで定義される。

【0066】図3は、ある1つのルータが、他のルータに対してハローパケットを送信したが、応答がなく、当該他のルータに障害が発生したと判断した場合に起動されるプロセスのフローである。当該プロセスは、図2で説明したプロセス同様、これ以外のタイミングでも起動されうる。

【0067】ステップS30において、次にハローパケットを正常に受信する場合に備えて、CPU使用率の測定を開始する。この場合、他のルータは障害中であって、ハローパケットを正常に送信できるようになるまでは、多くの時間を要する場合がある。従って、その場合には、こうして得られたCPU使用率から、ハローパケットを受信する直前の一定時間TにおけるCPU使用率を適切に取り出す処理が必要となる。

【0068】ステップS32では、自ルータと他のルータとの間で、 $PRI_2 = PRI_2(B)$ -

ここで、 $PRI_2(B)$ （基本プライオリティ）は、予め管理者又は利用者が設定する値であり、 $PRI_2(H)$ （補正プライオリティ）は、パケット数Yを元に

$$PRI_2(H) = K_2 * Y$$

但し、 K_2 は所定の係数である。Yは、ある一定時間Tの間におけるパケット数である。但し、Yとして、パケット数を表す他の指標、例えば、時間Tを複数のタイム

スロットに分割した場合の、各タイムスロットにおけるパケット数の最大値、最小値、あるいは平均値等を使用することもできる。

【0069】このフローを各ルータが実行することによって、障害の発生したルータが、それまで担当していた仮想ルータIDについてのパケットの処理を、いくつかの他のルータで自動的に引き継ぐことができる。この機能は、従来のVRRPによって実現されるものと同様である。

【0070】また、障害の発生したルータが復旧した場合、その時点でルータは、どのパケットの処理も担当していないので、非常に低いCPU使用率となっていることが予想される。しかし、その後、何回か他のルータとハローパケットの交換を行うことによって、いくつかの仮想ルータIDのパケットに関して、処理の担当をすることになり、次第にCPU使用率も上昇することになる。こうしてCPU使用率が他のルータと同程度になると、処理する仮想ルータIDはほぼ安定する。

【0071】この点、従来のVRRPでは、固定値であるプライオリティを単に比較するだけで、どの仮想ルータIDのパケットを担当するかが決められていたが、本発明の第1の実施形態のルータでは、新たな仮想ルータIDのパケットを処理することによるCPU使用率の上昇の程度が加味され、他のルータのCPU使用率とバランスをとるように、自動的に調整される。

【0072】次に、本発明の第2の実施形態のルータについて、図4を参照して説明する。

【0073】図4のルータ30は、パケット測定手段31、及びプライオリティ設定手段32を有している。

【0074】パケット測定手段31は、自身のルータで処理されたパケットの数（トライフィック）をカウントし、その数を値Yで返す機能を有する。ここでのパケット数は、特定のIPアドレスを宛先としたパケットや、特定のアプリケーションのためのパケットのみを対象とした数とすることも可能である。

【0075】プライオリティ設定手段32は、パケット測定手段31で測定されたパケット数Yを元にして、動的にプライオリティを設定する機能を有する。プライオリティPRI2は、以下の式3によって求められる。

【0076】 $PRI_2(H) \dots \text{ (式3)}$

して計算された値であり、以下の式4によって求められる。

【0077】 $\dots \text{ (式4)}$

スロットに分割した場合の、各タイムスロットにおけるパケット数の最大値、最小値、あるいは平均値等を使用することもできる。

【0078】また更に、基本プライオリティPRI2(B)、係数K2は、ルータ毎に、そのルータの性能等に応じて設定することができ、また、1つのルータの中でも、仮想ルータID毎に異なる値を設定することができる。

【0079】これによって、所定の単位時間当たりに処理するパケット数が多いルータほどプライオリティが低くなり、自動的に、処理パケット数の少ないルータにパケット処理が動的にシフトされる。また、上記式3、及び式4は、パケット数の多いものほどプライオリティが小さくなるように調整された式の一例を示したに過ぎず、本発明の範囲がこれに限られるものではない。当業者であれば、多くの他の類似のパターンを導出することができる。

【0080】本発明実施形態のルータは、CPU使用率を測定する第1の実施形態と比較して、パケット数のカウントのみによってルータの負荷を測定するものであるため、場合によっては、ルータに余分な負荷を与えること

$$PRI3 = PRI3(B) - \dots \quad (式5)$$

ここで、PRI3(B)は、予め管理者又は利用者が設定する基本プライオリティであり、PRI1(H)は、式2により求められた、CPU利用率による補正プライオリティ、PRI2(H)は、式4により求められた、パケット数による補正プライオリティである。

【0085】当該実施形態においては、ルータのCPU使用率の増加、及びパケット数の増加が、そのルータのプライオリティを下げるようになると式を設定することが必要である。従って、前記式5はそのような式の一例に過ぎず、当業者は、他の多くのバリエーションを容易に導出することができる。例えば、PRI1(H)とPRI2(H)にそれぞれ一定比率を乗じて、これらがPRI3に及ぼす影響を調整する等の方法が考えられる。

【0086】ここまで、本発明の負荷分散方法について、IPの一種であるVRRPに適用する例について説明してきた。しかし、ルータの負荷を何らかの方法で定量化してルータのプライオリティに反映させ、ルータ間の負荷を分散させる本発明の負荷分散方法は、ルータを利用した他のプロトコル、あるいは通信方法にも十分適用できるものである。従って、本発明は、VRRPに適用する場合に限定される必要はなく、プライオリティによってルータにおけるパケット処理の配分を定める全ての方法に適用可能である。また、本発明の方法については、ルータによって実現される例を示してきたが、例えば、ゲートウェイ等の他のネットワーク中継装置に応用することも可能である。

【0087】最後に、本発明の各実施形態のルータの構成について、図5を参照して説明する。第1ないし第3の実施形態のルータについて説明してきたが、ルータの

好ましい場合がある。

【0081】次に、本発明の第3の実施形態のルータについて説明する。このルータは、第1の実施形態のルータと第2の実施形態のルータとを組み合わせたものである。即ち、各ルータのプライオリティが、そのルータのCPU使用率、及び処理パケット数を元に計算されるものである。

【0082】通常、CPU使用率は、処理パケット数が増加すれば高くなるが、パケットの宛先、パケット長、パケットの受信タイミング等によっては、必ずしもそうならない場合がある。そのため、CPU使用率とパケット数の両方を、プライオリティ計算のファクターとすることによって、ルータの負荷を総合的にとらえようとするものである。

【0083】この場合のプライオリティPRI3を求める式5は、例えば以下のようなものである。

【0084】

PRI1(H) - PRI2(H)

物理的構成は基本的に同じである。

【0085】図5のルータ50は、CPU51、メモリ52、外部記憶装置53、入力装置54、出力装置55、記録媒体駆動装置56、ネットワーク・インターフェース57、及びこれらの構成要素51～57を相互に接続するバス58からなる。

【0086】CPU51は、ルータ50内の各構成要素の制御、及びルーティング処理、その他パケットの処理を実行するのに使用される。

【0087】メモリ52は、前記制御、及び各処理を行うようCPU51を制御するプログラム、及び当該プログラムが使用するデータを一時的に記憶する。

【0088】外部記憶装置53は、ルーティング・テーブルや、VRRPで使用される仮想ルータID(VRID)等を記憶し、更に、本発明の負荷分散方法を実行するのに必要な基本プライオリティ(PRI1(B)、PRI2(B)、PRI3(B))や、係数K1、K2などを記憶するハードディスク等の記憶装置である。

【0089】入力装置54は、マウス、キーボード等の、CPU51に特定の指令を与えたり、外部記憶装置53に記憶するデータを入力するための装置であり、上記仮想ルータID、基本プライオリティ、及び係数等は、管理者等によって当該入力装置54を介して入力される。

【0090】出力装置55は、CRTディスプレイやテレビモニタであるが、ルータの切り替えが生じた場合や、自ルータ、又は他のルータの障害を検知した場合にメッセージが表示される。

【0091】記録媒体駆動装置56は、DVD-ROM、CD-ROM、フレキシブルディスク等の記録媒体

59の内容を読み出し、必要に応じて前記メモリ52や外部記憶装置53にコピーする。本発明の負荷分散方法を実行するためにCPU51に指令を与え制御する前記プログラムも当該記録媒体59に格納される場合があり、必要に応じてメモリ52にロードされる。ルータ50は、このメモリ52内にロードされた当該プログラムによって制御され、CPU51はこのプログラムの命令に従って各種処理を実行し、本発明の方法を実現する。

【0095】ネットワーク・インターフェース57は、ルータ50をネットワーク60、又は他のネットワーク機器61に接続する。ここでは、他のネットワーク機器61は、例えば、図6に示されるようなホストである。ホストから送出されたパケットは、ネットワーク・インターフェース57を介してルータ50に送信され、ルータ50によって、送信経路が判断され、適当な別なルータ、或いはネットワークに、ネットワーク60を介して送信される。

【0096】前記CPU51に指令を与え、本発明の方法を実現するプログラムは、当該ネットワーク60及びネットワーク・インターフェース57を介してメモリ52にロードされる場合もある。

【0097】

【発明の効果】本発明のルータによって、VRRPを使用した場合に、各ルータのブレイブリティを、そのルータの負荷に応じて調整し、ルータ間の負荷を分散させることができる。

【0098】のことにより、従来、VRRPでは、マスター・ルータが完全に通信不能にならないと前述のルータの切替が行われなかつたが、マスター・ルータにトラフィックが集中して、パケットのルーティング処理等の処理効率が著しく低下している場合であつても、そのルータの負荷を判定して、パケットの処理を、自動的に他のバックアップ・ルータに切り替えることができる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態のルータの機能ブロック図である。

【図2】本発明の第1の実施形態のルータにおいて、ハローパケットを受信した際の処理を示すフローチャートである。

【図3】本発明の第1の実施形態のルータにおいて、他のルータの障害を検知した際の処理を示すフローチャートである。

【図4】本発明の第2の実施形態のルータの機能ブロック図である。

【図5】本発明のルータの構成を示すブロック図である。

【図6】従来のVRRPの機能を説明するための、ルータ及びホストを含むネットワーク構成図である。

【符号の説明】

10、30、50、71、72 ルータ

11 CPU使用率測定手段

12、32 ブライオリティ設定手段

31 パケット測定手段

51 CPU

52 メモリ

53 外部記憶装置

54 入力装置

55 出力装置

56 記録媒体駆動装置

57 ネットワーク・インターフェース

58 バス

59 記録媒体

60 ネットワーク

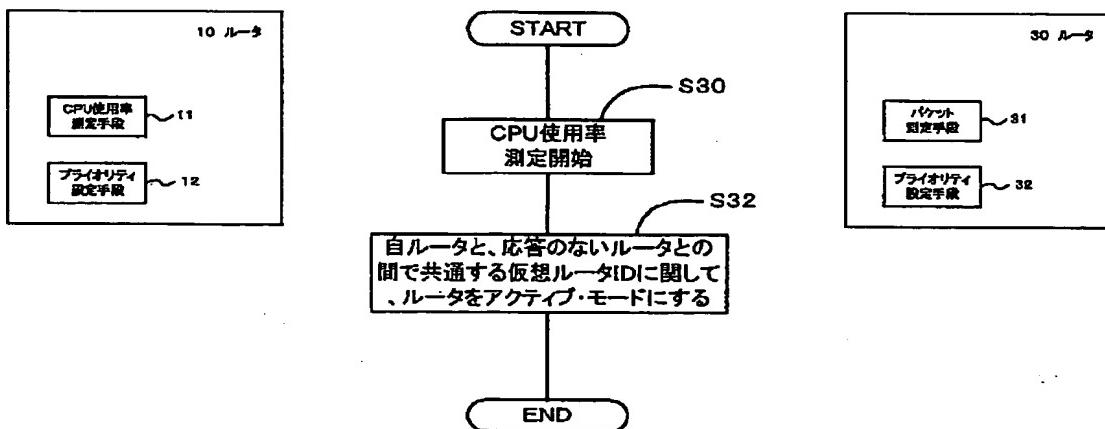
61 ネットワーク機器

81、82、83、84 ホスト

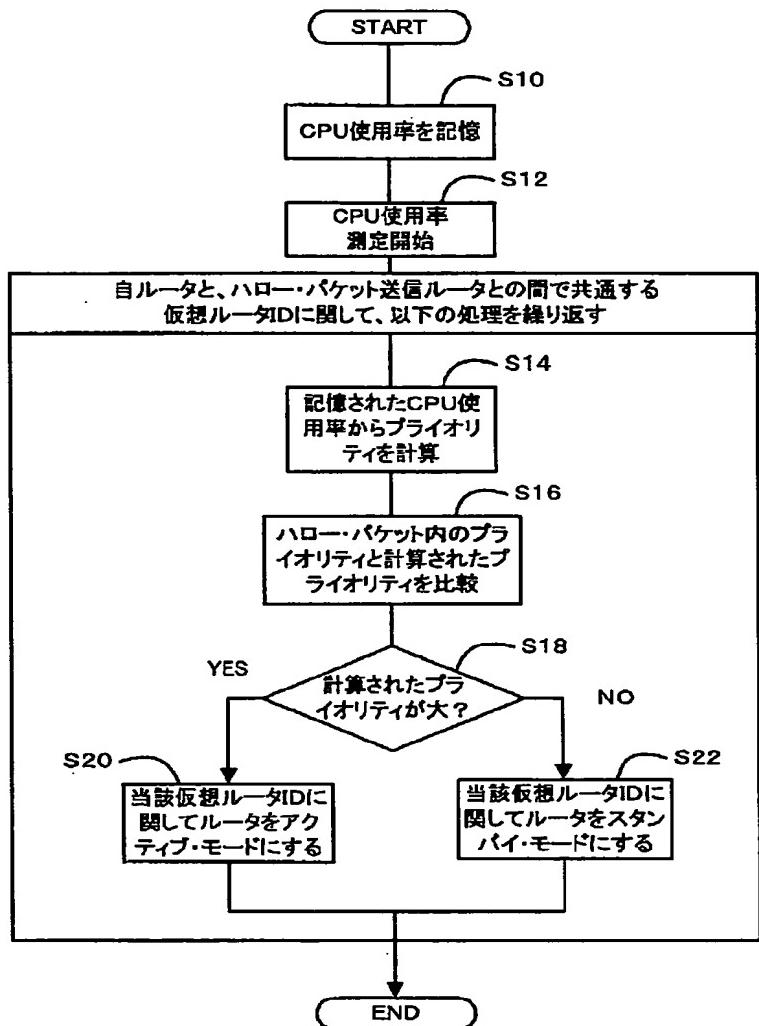
【図1】

【図3】

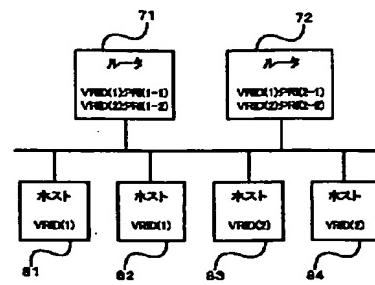
【図4】



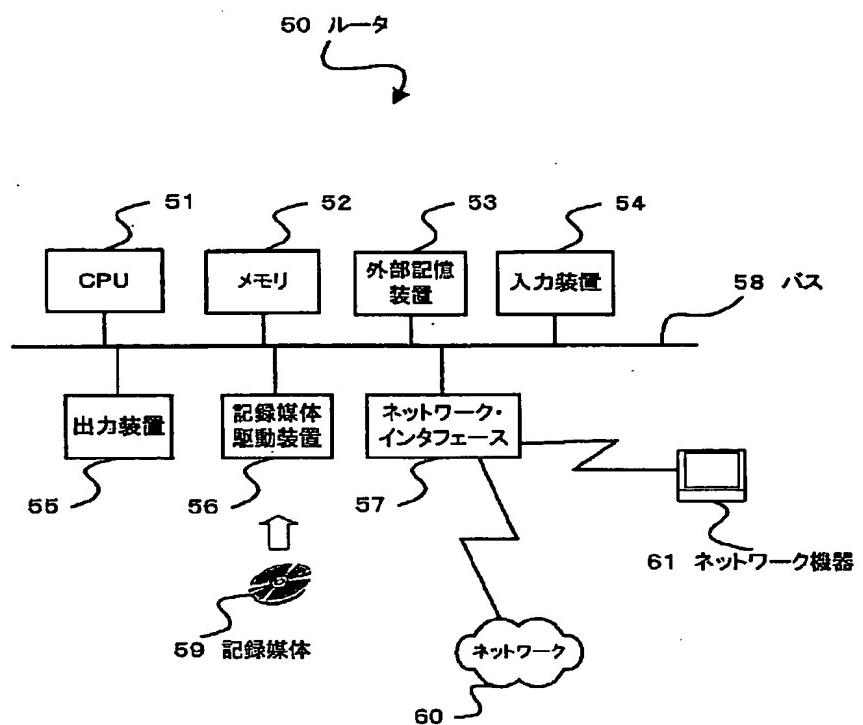
【図2】



【図6】



【図5】



This Page Blank (uspto)

This Page is inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT OR DRAWING
- BLURED OR ILLEGIBLE TEXT OR DRAWING
- SKEWED/SLANTED IMAGES
- COLORED OR BLACK AND WHITE PHOTOGRAPHS
- GRAY SCALE DOCUMENTS
- LINES OR MARKS ON ORIGINAL DOCUMENT
- REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning documents *will not* correct images problems checked, please do not report the problems to the IFW Image Problem Mailbox

This Page Blank (uspto)